Journal of Educational Psychology
2000, Vol. 92, No. 4, 718–723

0022-0663/00/$5.00 DOI: 10.1037//0022-0663.92.4.718

Gender Differences in Above-Level EXPLORE Scores of Gifted Third Through Sixth Graders

Mary Ann Swiatek
Lafayette College

Ann Lupkowski-Shoplik
Carnegie Mellon University

Cathleen C. O'Donoghue
The New England Center for Children

We examined gender differences in EXPLORE scores when taken by gifted 3rd through 6th graders. Boys performed better on Mathematics and Science Reasoning, and girls performed better on Reading, but effect sizes were negligible. In English, boys scored higher in third grade, and girls scored higher in subsequent grades. More boys than girls scored at or above a cutoff of 14 on Mathematics, and more girls than boys scored at this level on English and Reading. Using a cutoff of 25, the male advantage in Mathematics and Science Reasoning increased, but there was no gender difference in English or Reading. These findings parallel those from studies of gifted seventh and eighth graders: Test performance of boys in Mathematics was somewhat stronger than that of girls, regardless of how performance was measured, but results favoring girls in verbal areas were weaker and less consistent.

The talent search model uses above-level standardized tests to determine the extent of gifted students' academic abilities. This model, pioneered in the early 1970s by Julian C. Stanley at Johns Hopkins University to serve gifted students at the junior high school level, identifies as potential participants those students who score in the top 3 to 5% on in-grade standardized achievement tests. Seventh and eighth graders who do so are given the opportunity to take a college entrance exam (i.e., the Scholastic Aptitude Test [SAT] or the ACT Assessment) as an above-level test (e.g., George, 1979; Keating, 1974; Olszewski-Kubilius, 1998; Stanley, 1988; VanTassel-Baska, 1984). Using a test that is designed for individuals several years older serves to raise the score ceiling above that of in-grade tests, thereby "spread[ing] out the scores of able students, helping to differentiate talented students from the exceptionally talented students" (Rotigel & Lupkowski-Shoplik, 1999, p. 331; see also Colangelo, Assouline, & Lu, 1994; George, 1979; Olszewski-Kubilius, 1998; VanTassel-Baska, 1984). Further, because students usually have not been formally exposed to the material covered in above-level tests, the tests are viewed as measures of reasoning ability rather than academic achievement (Stanley & Benbow, 1986). Talent searches are currently available to junior-high level students throughout the United States through regional talent search programs at Johns Hopkins University, Duke University, Northwestern University, and the University of Denver. In addition, many state talent search centers exist (Assouline & Lupkowski-Shoplik, 1997; Olszewski-Kubilius, 1998).

Recently, the talent search model was extended to serve elementary school students (Colangelo et al., 1994). The EXPLORE test, which was developed to measure eighth graders' curriculum-related knowledge and complex cognitive skills (ACT, 1997), is used as an above-level instrument in talent searches sponsored by Carnegie Mellon University, Duke University, Northwestern University, and the University of Iowa (Assouline & Lupkowski-Shoplik, 1997). EXPLORE, which was first used as an above-level test for gifted elementary students in 1993, provides a ceiling high enough to distinguish among different levels of high ability in third through sixth graders (ACT, 1996). Scores earned by talent search participants span the full range of possible test scores (Assouline & Lupkowski-Shoplik, 1997; Colangelo et al., 1994; Lupkowski-Shoplik & Swiatek, 1999). EXPLORE consists of four 30-min multiple-choice subtests: English, Mathematics, Reading, and Science Reasoning. The test taker receives a score for each subtest and a composite score that represents the average across subtests (ACT, 1997).

Because elementary student talent searches are relatively new, little research pertaining to the performance of participants has been conducted. In contrast, many reports have been published about the test performance of junior-high level talent search students. Much of this research was conducted by the Study of Mathematically Precocious Youth (SMPY), a longitudinal study of talent search participants, which like the talent search itself, was founded by Stanley (see Lubinski & Benbow, 1994). One of the

Mary Ann Swiatek, Department of Psychology, Lafayette College; Ann Lupkowski-Shoplik, Carnegie Mellon Institute for Talented Elementary Students, Carnegie Mellon University; Cathleen C. O'Donoghue, The New England Center for Children, Southboro, Massachusetts.

Mary Ann Swiatek is now at the Carnegie Mellon Institute for Talented Elementary Students, Carnegie Mellon University.

Correspondence concerning this article should be addressed to Mary Ann Swiatek, Carnegie Mellon Institute for Talented Elementary Students, 4902 Forbes Avenue, Box 6261, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213. Electronic mail may be sent to swiatekm@andrew.cmu.edu.

findings to emerge from SMPY research is that, in large samples of high ability seventh-grade students, boys consistently obtained higher mean scores than did girls on the mathematics section of the SAT (i.e., SAT-M; Benbow & Stanley, 1980, 1983). Further, the proportion of boys to girls achieving a given SAT-M score rose as the score being considered rose. About twice as many boys as girls scored at or above 500 on the SAT-M (e.g., Benbow, 1988; Benbow & Stanley, 1980, 1983; Goldstein & Stocking, 1996; Stanley, 1988, 1993), whereas about 12 times as many boys as girls scored 700 or more (Benbow, 1988; Benbow & Stanley, 1983; Stanley, 1988). Some authors (e.g., Brody, Barnett, & Mills, 1996; Goldstein & Stocking, 1996) reported more variable gender ratios at the 700 cutoff level, but these ratios were consistently higher than those reported for scores at or above 500.

Recently, Robinson and her colleagues studied mathematical reasoning in young children, starting at the preschool and kindergarten levels (Robinson, Abbott, Berninger, & Busse, 1996) and following up over a 2-year period (Robinson, Abbott, Berninger, Busse, & Mukhopadhyah, 1997). Participants were those children nominated by their parents as having advanced reasoning or interest in mathematics and who subsequently scored at the 98th percentile or higher on the mathematics subtest of at least one of three standardized cognitive tests. Data collection involved the administration of a battery of additional cognitive measures.

Robinson et al. (1996) found a number of gender differences, favoring boys, in their final sample:

> More boys than girls were nominated [for the study]; of those nominated, more boys than girls qualified; and, on the psychometric battery, boys' scores in the mathematical domain were significantly higher than the girls' on 8 of the 11 mathematical subtests. (p. 350)

They also found a significantly greater proportion of boys than girls in the top 5% of the group on a number of the mathematics subtests. Further, over the course of the next 2 years, the boys gained more ground in mathematics than did the girls (Robinson et al., 1997).

Because above-level testing for gifted elementary students is now available, it is possible to investigate gender differences in mathematical reasoning among individuals in third through sixth grade. Gender differences favoring boys have been documented, but the magnitude of these differences has not been consistent. Colangelo et al. (1994) and Assouline and Doellinger (1999) reported significantly higher scores among boys than girls on the above-level EXPLORE Mathematics subtest. We calculated the effect size (d) of this gender difference to be .17 in each article (see Glass, McGaw, & Smith, 1981, for formula), which is negligible by Cohen's (1988) standards. Mills, Ablard, and Stumpf (1993) used the School and College Ability Test (SCAT) as an above-level instrument to examine gender differences in mathematical reasoning among gifted second through sixth graders. They found differences in mean scores, favoring boys, on the mathematics portion of the SCAT at every grade level studied. Effect sizes for these differences ranged from .40 to .50 (in the small to medium range). Stanley (1994) examined the Secondary School Admission Test—Upper Level (SSAT-U), which was intended for eighth to tenth graders, as an above-level instrument for gifted fifth and sixth graders. On the quantitative subtest, he found differences favoring boys at both grade levels; these differences were characterized by small effect sizes (i.e., $d = .36$ for fifth graders and $d = .44$ for sixth graders). Small values of d may accompany large gender differences in the upper tail of the distribution (Stanley, 1994), however, even if the groups have equal variances (Feingold, 1995). For example, Stanley's (1994) small effect sizes were obtained from a sample in which approximately seven times more boys than girls achieved scores in the top 1% of the distribution.

No gender differences were found in mean above-level test scores on the verbal section of the SAT (SAT-V) when it was used with gifted seventh and eighth graders (e.g., Brody et al., 1996; Stanley, 1994; Weiner & Robinson, 1986) or on the SSAT-U when it was used with gifted fifth and sixth graders (Brody et al., 1996). In 1994, Colangelo et al. reported a significant difference favoring girls in mean EXPLORE reading scores among gifted third through sixth graders. This difference appeared to be counter to the findings in other studies, but its effect size was negligible in magnitude. The only nonnegligible effect size obtained by Colangelo et al. (1994) was for a mean difference favoring girls in English; this effect size was small. The interaction between gender and grade level was significant for the English subtest, indicating that the girls' advantage increased from Grade 3 to Grade 6. In contrast to the lack of gender differences found in these studies, a modest advantage for boys was often found by researchers who compared proportions of boys and girls scoring in the upper tail of an above-level test distribution. This advantage also was documented using the SAT-V (Goldstein & Stocking, 1996) and the SSAT-U (Stanley, 1994).

Overall, when means were compared, there were no gender differences in above-level EXPLORE scores. The ratio of boys to girls in the upper tail of the EXPLORE distribution has not been investigated, however, despite the fact that differences in the upper tail can be meaningful even when differences between means are not (Feingold, 1995; Hedges & Nowell, 1995; Stanley, 1994; Stanley, Benbow, Brody, Dauber, & Lupkowski, 1992). In the present study, we used two approaches to determine whether there were gender differences on any of the four subtests of EXPLORE when it was given to gifted third through sixth graders as an above-level instrument. First, mean scores were compared to address the following research questions: (a) Did boys and girls differ in the mean scores they earned on each EXPLORE subtest? and (b) Did any gender differences change with age? Second, proportions of boys and girls earning high scores were compared to address two more questions: (a) Were the genders evenly represented in the upper tails of the score distributions? and (b) Did the ratio of boys to girls change as cutoff scores were raised?

Method

Participants

Participants were the 5,422 third- through sixth-grade students (2,949 boys, 2,471 girls, and 2 individuals who did not report their gender) who, from 1997 through 1999, participated for the first time in the Elementary Student Talent Search conducted by the Carnegie Mellon Institute for Talented Elementary Students (C-MITES). A score at the 95th percentile or higher on the composite score or on the vocabulary, reading, math total, or science subtest of an in-grade standardized achievement test was required for students to qualify for the talent search (C-MITES, 1997, 1998, 1999).

Ninety-nine percent of the participants were from the Commonwealth of Pennsylvania. The majority of participants (approximately 85%) were

Caucasian American/White. Other members of the sample reported their ethnicity as follows: African American/Black (1.6%), American Indian/Alaskan Native (0.1%), Mexican American/Chicano (0.1%), Asian American/Pacific Islander (3.8%), Puerto Rican/Cuban/Other Hispanic (0.2%), Multiracial (2.2%), Other (1.0 %); 1.1% of the sample left the ethnicity item blank, and 1.0% marked an option stating "I prefer not to respond." These figures are quite similar to those found in the 1990 Pennsylvania census (U.S. Bureau of the Census, 1990), except that African American/Black students are underrepresented (the Pennsylvania census reports 9.2%) and Asian American/Pacific Islander students are overrepresented (the Pennsylvania census reports 1.2%). Other talent search programs have reported similar demographics, with underrepresentation of African Americans and overrepresentation of Asian Americans (see Ebmeier & Schmulbach, 1989; Granovetter & Olszewski-Kubilius, 1992).

Five hundred fifty-nine individuals (329 boys and 230 girls) were in the third grade, 2,508 (1,337 boys and 1,171 girls) were in the fourth grade, 1,792 (986 boys and 806 girls) were in the fifth grade, and 562 (297 boys and 265 girls) were in the sixth grade when EXPLORE was administered. There were relatively few third graders in the participant group because the C-MITES staff encouraged parents to exercise caution in deciding whether to register third graders for the talent search. It was suggested (although not required) that third graders participate only if they scored at or above the 95th percentile on all qualifying subtests of an in-grade achievement test, instead of only one. There were relatively few sixth graders in the participant group because sixth graders were not invited to participate in the C-MITES Elementary Student Talent Search until 1998.

Instrumentation

Information regarding participants' gender and grade level was taken from the talent search consent/registration form. Academic reasoning ability was measured with EXPLORE, a multiple-choice test administered by ACT. The test was designed for eighth-grade students to measure development in English (40 items), Mathematics (30 items), Reading (30 items), and Science Reasoning (28 items). ACT provided a raw score (i.e., the number of items correct), a scaled score (ranging from 1–25), and a percentile rank (ranging from 1–99 and based on the performance of eighth graders) for each subtest and for the composite, which represents the average performance across subtests.

Procedure

Students were invited to participate in the talent search on the basis of their scores on a variety of in-grade achievement tests, including the Iowa Test of Basic Skills, the California Achievement Test, the Stanford Achievement Test, the Comprehensive Test of Basic Skills, and the Metropolitan Achievement Test. Information about the C-MITES Elementary Student Talent Search was mailed to all schools in the Commonwealth of Pennsylvania, and school personnel were asked to search their standardized test score records and notify all eligible students about the C-MITES program. These personnel were informed that students may be involved in the C-MITES Elementary Student Talent Search regardless of their ability to pay; students participating in the free or reduced-cost lunch program at school automatically received a fee waiver for testing. Separate mailings were sent to gifted coordinators and principals, and a reminder mailing was sent several months later to gifted coordinators. The majority of students learned about the C-MITES program through their schools, although others learned about it through newspaper articles, newsletter ads, flyers posted in local libraries, and word of mouth.

Registration packets were mailed each September, and EXPLORE was administered the following January and February. A registration fee of $49 to $54 was assessed (except as noted earlier) to cover testing, handouts, and a 3-year subscription to the C-MITES semiannual newsletter. Participants took EXPLORE with other elementary students at a number of talent search test centers throughout Pennsylvania. ACT scored the tests and provided score reports to C-MITES and to the students' families.

A multivariate analysis of variance (MANOVA) was conducted with gender and grade level as independent variables and EXPLORE subtest scores as dependent variables. Chi-square analyses also were performed. First, an EXPLORE subtest cutoff of 14 was used; this score approximates average performance in the eighth-grade norm group on each subtest. Second, similar analyses were performed with a cutoff reflecting approximately the 95th percentile of the eighth-grade norm group (scaled scores were 22 for Science Reasoning, 22 for Mathematics, 21 for English, and 23 for Reading; ACT, 1997).

Results

The MANOVA yielded significant main effects for gender, $F(4, 5409) = 30.97, p < .001$, grade level, $F(12, 16233) = 159.00, p < .001$, and for their interaction, $F(12, 16233) = 1.94$, $p < .05$. Univariate tests indicated significant gender differences on all four EXPLORE subtests. Boys scored higher than girls on the Mathematics subtest (M = 13.22 vs. 12.62), $F(1, 5412) = 34.50$, $p < .001$, $d = .18$, and the Science Reasoning subtest (M = 15.02 vs. 14.69), $F(1, 5412) = 11.70, p < .005, d = .08$, and girls scored higher than boys on the Reading subtest (M = 14.07 vs. 13.51), $F(1, 5412) = 13.64, p < .001, d = -.12$, and the English subtest (M = 15.02 vs. 14.58), $F(1, 5412) = 6.32, p < .05, d = -.11$, but all effect sizes were negligible. Grade level differences also were statistically significant for all four subtests. Scheffe post hoc tests indicated significant increases in all four subtest scores from each grade to the next. A significant interaction between gender and grade was found only for the English subtest, $F(3, 5412) = 4.27$, $p < .01$.

The results of chi-square (χ^2) analyses are presented in Table 1. Alpha levels were adjusted to accommodate the multiple comparisons; instead of using $p < .05$ as the criterion for statistical significance, each set of chi-square analyses was based on a value of $p < .013$ (i.e., .05/4), which was selected using Bonferroni's approach (see Howell, 1987).

There was no significant gender difference in the proportion of boys versus girls scoring 14 or higher on the Science Reasoning subtest. More girls than boys scored at this level on the Reading subtest (54% vs. 49%), $\chi^2(1, N = 5{,}420) = 12.00$, $p < .005$, and the English subtest (67% vs. 61%), $\chi^2(1, N = 5{,}420) = 19.61$, $p < .001$, whereas more boys than girls achieved this score on the Mathematics subtest (42% vs. 35%), $\chi^2(1, N = 5{,}420) = 27.17$, $p < .001$.

When EXPLORE cutoffs were established to reflect approximately the 95th percentile of the eighth-grade norms, no significant gender differences were found for the English or Reading subtests. More boys than girls achieved scores at this level on the Mathematics subtest (3% vs. 1%), $\chi^2(1, N = 5{,}420) = 13.79$, $p < .001$, and the Science Reasoning subtest (6% vs. 3%), $\chi^2(1, N = 5{,}420) = 17.89$, $p < .001$.

Discussion

It was not surprising that there was a steady and significant increase in scores on all EXPLORE subtests as grade level in-

Table 1
Gender Comparisons at Two EXPLORE Cutoff Levels

Cutoff level and subtest	% boys	% girls	Ratio (boys:girls)	$\chi^2(1, N = 5{,}420)$
Cutoff level at 14[a]				
Mathematics	42.2	35.2	1.20:1	27.17***
Science Reasoning	62.3	60.2	1.04:1	2.45
Reading	49.1	53.8	0.91:1	12.00**
English	61.2	67.0	0.91:1	19.61***
Cutoff level at 95th percentile[b]				
Mathematics	2.5	1.1	2.27:1	13.79***
Science Reasoning	5.9	3.4	1.74:1	17.89***
Reading	3.0	3.2	0.94:1	0.28
English	5.7	7.0	0.81:1	4.34

[a] Refers to students scoring 14 or higher on a given subtest. [b] Refers to students scoring at or above the 95th percentile compared with the eighth-grade norm group.
** $p < .005$. *** $p < .001$.

creased. Of greater interest were the main effects for gender and the interaction between gender and grade level. It was hypothesized that boys would score higher than girls on the Mathematics subtest of EXPLORE (see Assouline & Doellinger, 1999; Benbow & Stanley, 1980, 1983; Brody et al., 1996; Colangelo et al., 1994; Mills et al., 1993; Robinson et al., 1996, 1997; Stanley, 1994). This hypothesis was supported, but the negligible effect size suggests that the power of the large sample size may have accounted for the statistical significance. This finding is similar to those of previous studies of gifted elementary school students, in which gender differences in mathematics scores were statistically significant and favored boys, but effect sizes were negligible (Assouline & Doellinger, 1999; Colangelo et al., 1994) or small (Mills et al., 1993).

Similar findings were obtained for the remaining three EXPLORE subtests. In each case, a significant gender difference was found, but the effect size was negligible. Although gender differences in verbal areas were sometimes found among seventh and eighth graders taking above-level tests (Goldstein & Stocking, 1996), these differences were small, and other researchers have failed to find any gender differences at all (Brody et al., 1996; Weiner & Robinson, 1986). Findings for elementary school students have been somewhat more consistent; there were usually no meaningful gender differences in mean above-level test scores in the verbal area (Brody et al., 1996; Colangelo et al., 1994; Stanley, 1994). Our results are consistent with these findings.

Our results are also similar to those obtained by Colangelo et al. (1994) in that the only significant interaction between gender and grade level was found for the English subtest. Our third-grade boys scored slightly higher than did third-grade girls, but the advantage was reversed among fourth graders, and girls remained ahead through the sixth grade. The interpretation of this interaction is limited by the fact that the data are cross-sectional, but the interaction replicated that found by Colangelo et al. (1994) in a different research sample. Therefore, gifted girls' improvement in written English may outpace that of gifted boys during the elementary school years.

The effect sizes of the gender differences in mean EXPLORE scores do not present a complete picture of gender differences in tested ability, however. Even when effect sizes for mean differences are negligible, differences in the proportions of boys and girls in the upper tail of the score distributions can be striking (Feingold, 1995; Hedges & Nowell, 1995; Stanley, 1994; Stanley et al., 1992). We examined the proportion of boys and girls scoring at or above two cutoff levels on the various subtests of EXPLORE. The first cutoff was 14, which was chosen using the same rationale as the SAT-M cutoff of 500 that was used to analyze data from older talent search participants (e.g., Benbow, 1988; Benbow & Stanley, 1980, 1983; Goldstein & Stocking, 1996; Stanley, 1988, 1993). Both scores approximate the average score among members of the test's norm group (i.e., eighth graders for EXPLORE and twelfth graders for the SAT). Boys were significantly more likely than girls to score at or above 14 on EXPLORE Mathematics, with a ratio of 1.2:1, whereas girls were significantly more likely than boys to reach this cutoff on EXPLORE Reading and English, with ratios of 1.1:1 for both subtests. On the Science Reasoning subtest, boys and girls were equally likely to score above the cutoff of 14. Although these ratios are rather low, interesting changes occurred when EXPLORE cutoffs were raised.

At the higher cutoff (the 95th percentile when compared with the normative group), significant gender differences favoring girls in English and reading were eliminated. A significant difference favoring boys appeared for science reasoning; there were 1.74 boys for every 1 girl at this cutoff level. The significant difference favoring boys in math also increased to a ratio of 2.27:1. Although not as extreme as the findings for seventh-grade talent search students, who have shown ratios as high as 6:1 or even 13:1 at the highest scores in mathematics (e.g., Benbow, 1988; Benbow & Stanley, 1980, 1983; Goldstein & Stocking, 1996; Stanley, 1988, 1993), the same pattern was evident. Further, this pattern could not be explained by grade-level differences among the students who reached or surpassed the cutoff scores. Where statistically significant differences in grade level existed, they were very small. Girls in the upper tail of the distribution had a slightly higher mean grade level than did boys (i.e., grade level M = 4.95 for girls and 4.80 for boys earning an EXPLORE Mathematics score at or above 14; M = 5.75 for girls and 5.09 for boys earning a score at or above the 95th percentile).

Consistent with previous work (Mills et al., 1993; Robinson et al., 1996), gender differences in test performance were found well before junior high school among gifted students in this sample. Although such gender differences were not apparent in mean EXPLORE scores among elementary school students, they were more obvious in the ratio of boys to girls among high mathematics scorers. Gender differences favoring girls existed in reading and English at moderately high cutoff levels (i.e., the average score from the eighth-grade norms), but these differences were modest and disappeared at very high cutoff levels (i.e., the 95th percentile from the eighth-grade norms). In contrast, gender differences favoring boys in math increased at high cutoff levels. Although these differences were relatively small in third through sixth graders, previous researchers (Benbow, 1988; Benbow & Stanley, 1980, 1983; Goldstein & Stocking, 1996; Stanley, 1988, 1993) have shown that they are larger among seventh and eighth graders. Considering the pattern of gender differences evidenced in this research, it appears important to provide special programming for mathematically and scientifically gifted girls during elementary school.

References

ACT, Inc. (1996). *Evaluating the appropriateness of EXPLORE for BESTS (Belin Elementary Student Talent Search) participants.* Iowa City, IA: Author.

ACT, Inc. (1997). *EXPLORE* technical manual. Iowa City, IA: Author.

Assouline, S. G., & Doellinger, H. L. (1999). *Evidence of the underchallenging elementary mathematics curriculum: Implications for instruction.* Manuscript submitted for publication.

Assouline, S. G., & Lupkowski-Shoplik, A. (1997). Talent searches: A model for the discovery and development of academic talent. In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (2nd ed., pp. 170–179). Boston: Allyn and Bacon.

Benbow, C. P. (1988). Sex differences in mathematical reasoning ability in intellectually talented preadolescents: Their nature, effects, and possible causes. *Behavioral and Brain Sciences, 11,* 169–232.

Benbow, C. P., & Stanley, J. C. (1980, December). Sex differences in mathematical ability: Fact or artifact? *Science, 210,* 1262–1264.

Benbow, C. P., & Stanley, J. C. (1983, December). Sex differences in mathematical reasoning ability: More facts. *Science, 222,* 1029–1031.

Brody, L. E., Barnett, L. B., & Mills, C. J. (1996). Gender differences among talented adolescents: Research studies by SMPY and CTY at The Johns Hopkins University. In K. A. Heller & E. A. Hany (Eds.), *Competence and responsibility: Proceedings of the Third European Conference of the European Council for High Ability* (pp. 204–210). Goettingen, Germany: Hogrefe and Huber.

Carnegie Mellon Institute for Talented Elementary Students. (1997). *Carnegie Mellon 1997 Elementary Student Talent Search: Registration information.* Pittsburgh, PA: Author.

Carnegie Mellon Institute for Talented Elementary Students (1998). *Carnegie Mellon 1998 Elementary Student Talent Search: Registration information.* Pittsburgh, PA: Author.

Carnegie Mellon Institute for Talented Elementary Students (1999). *Carnegie Mellon 1999 Elementary Student Talent Search: Registration information.* Pittsburgh, PA: Author.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (Rev. ed.). San Diego, CA: Academic Press.

Colangelo, N., Assouline, S. G., & Lu, W. -H. (1994). Using EXPLORE as an above-level instrument in the search for elementary student talent. In N. Colangelo, S. G. Assouline, & D. L. Ambroson (Eds.), *Talent development: Proceedings from the 1993 Henry B. and Jocelyn Wallace National Research Symposium on Talent Development* (pp. 281–297). Dayton, OH: Ohio Psychology Press.

Ebmeier, H., & Schmulbach, S. (1989). An examination of the selection practices used in the Talent Search Programs. *Gifted Child Quarterly, 33,* 134–141.

Feingold, A. (1995). The additive effects of differences in central tendency and variability are important in comparisons between groups. *American Psychologist, 50,* 5–13.

George, W. C. (1979). The talent search concept: An identification strategy for the intellectually gifted. *The Journal of Special Education, 13,* 221–237.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research.* Beverly Hills, CA: Sage.

Goldstein, D., & Stocking, V. B. (1996). TIP studies of gender differences in talented adolescents. In K. A. Heller & E. A. Hany (Eds.), *Competence and responsibility: Proceedings of the Third European Conference of the European Council for High Ability* (pp. 190–203). Goettingen, Germany: Hogrefe and Huber.

Granovetter, E., & Olszewski-Kubilius, P. (1992). *Center for Talent Development 1992 Midwest Talent Search: Report on participants.* Evanston, IL: Center for Talent Development.

Hedges, L. V., & Nowell, A. (1995, October). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science, 269,* 41–45.

Howell, D. C. (1987). *Statistical methods for psychology* (2nd ed.). Boston: PWS-Kent.

Keating, D. P. (1974). The Study of Mathematically Precocious Youth. In J. C. Stanley, D. P. Keating, & L. H. Fox (Eds.), *Mathematical talent: Discovery, description, and development* (pp. 23–46). Baltimore: Johns Hopkins University Press.

Lubinski, D., & Benbow, C. P. (1994). The Study of Mathematically Precocious Youth: The first three decades of a planned 50-year study of intellectual talent. In R. F. Subotnik & K. D. Arnold (Eds.), *Beyond Terman: Contemporary longitudinal studies of giftedness and talent* (pp. 255–281). Norwood, NJ: Ablex.

Lupkowski-Shoplik, A., & Swiatek, M. A. (1999). Elementary student talent searches: Establishing appropriate guidelines for qualifying test scores. *Gifted Child Quarterly, 43,* 265–272.

Mills, C. J., Ablard, K. E., & Stumpf, H. (1993). Gender differences in academically talented young students' mathematical reasoning: Patterns across age and subskills. *Journal of Educational Psychology, 85,* 340–346.

Olszewski-Kubilius, P. (1998). Talent search: Purposes, rationale, and role in gifted education. *Journal of Secondary Gifted Education, 9,* 106–113.

Robinson, N. M., Abbot, R. D., Berninger, V. W., & Busse, J. (1996). The structure of abilities in mathematically precocious young children: Gender similarities and differences. *Journal of Educational Psychology, 88,* 341–352.

Robinson, N. M., Abbott, R. D., Berninger, V. W., Busse, J., & Mukhopadhyah, S. (1997). Developmental changes in mathematically precocious young children. *Gifted Child Quarterly, 41,* 145–158.

Rotigel, J. V., & Lupkowski-Shoplik, A. (1999). Using talent searches to identify and meet the educational needs of mathematically talented youngsters. *School Science and Mathematics, 99,* 330–337.

Stanley, J. C. (1988). Some characteristics of SMPY's "700–800 on SAT-M before age 13 group": Youths who reason extremely well mathematically. *Gifted Child Quarterly, 32,* 205–209.

Stanley, J. C. (1993). Boys and girls who reason well mathematically. In G. R. Bock & K. Ackrill (Eds.), *The origins and development of high ability* (pp. 119–138). New York: Wiley.

Stanley, J. C. (1994). Gender differences for able elementary school students on above-grade-level ability and achievement tests. In N.

Colangelo, S. G. Assouline, & D. L. Ambroson (Eds.), *Talent development: Proceedings from the 1993 Henry B. and Jocelyn Wallace National Research Symposium on Talent Development* (pp. 141–148). Dayton, OH: Ohio Psychology Press.

Stanley, J. C., & Benbow, C. P. (1986). Youths who reason exceptionally well mathematically. In R. J. Sternberg & J. E. Davidson (Eds.), *Conceptions of giftedness* (pp. 362–387). New York: Cambridge University Press.

Stanley, J. C., Benbow, C. P., Brody, L. E., Dauber, S., & Lupkowski, A. (1992). Gender differences on eighty-six nationally standardized aptitude and achievement tests. In N. Colangelo, S. G. Assouline, & D. L. Ambroson (Eds.), *Talent development: Proceedings from the 1991 Henry B. and Jocelyn Wallace National Research Symposium on Talent Development* (pp. 42–65). Unionville, NY: Trillium Press.

U. S. Bureau of the Census. (1990). *1990 U. S. census data.* Washington, DC: Author. Retrieved October 1, 1999, from the World Wide Web: http://venus.census.gov/cdrom/lookup/969050474

VanTassel-Baska, J. (1984). The talent search as an identification model. *Gifted Child Quarterly, 28,* 172–176.

Weiner, N. C., & Robinson, S. E. (1986). Cognitive abilities, personality, and gender differences in math achievement of gifted adolescents. *Gifted Child Quarterly, 30*(2), 83–87.

Received November 4, 1999
Revision received March 30, 2000
Accepted March 30, 2000 ■